

sEMG-BASED SPEECH UNIT IDENTIFICATION IN PHONETIC LANGUAGE USING MULTILAYER PERCEPTRON AND ENSEMBLE BOOSTING ALGORITHMS

Alfiya Parveen PC [1], Remya R Nair[1], Vasudev V[2], Venugopal G[1]

[1] Biomedical Instrumentation and Signal Processing, Department of Instrumentation and Control Engineering NSS College of Engineering, Palakkad, Kerala, 678008, India, (Affiliated to APJ Abdul Kalam Technological University, Kerala, India)

[2] Department of Electrical & Electronics Engineering, NSS College of Engineering, Palakkad, Kerala, 678008, India, (Affiliated to APJ Abdul Kalam Technological University, Kerala, India)

Corresponding Author: Alfiya Parveen P C
Email: alphy000@gmail.com
Doi: 10.34107/TGHN9103067

ABSTRACT

Language is the most important means of communication, where orthography is the method of writing a specific language. Phonetic languages exhibit an exact mapping between letters and their corresponding sounds. Whereas, in non-phonetic languages shows irregular mapping between graphemes and phonemes. The proposed study primarily investigates the recognition of phonetic sounds in Malayalam, using surface electromyography (sEMG). The facial and neck muscles taken for the study are zygomaticus major, mentalis, depressor anguli oris and the anterior belly of digastric. Signals are collected from 45 healthy female participants using a well-defined protocol. The extraction of features includes five time-domain features and forty frequency-domain features. This includes high and low frequency average energy (E_high & E_low), high and low-frequency average amplitude (MAV_low & MAV_hig), and high frequency zerocrossing rate (ZCR_high) in the time domain, and 40 mel energy filter bank features are taken in the frequency domain. Further, the features extracted are fed into the machine learning model, such as multilayer perceptron and ensemble boosting algorithms, including XG-Boost, CatBoost, and Gradient Boost, for comparative analysis. Among the trained models, the highest classification accuracy of 83% is obtained with the XG-Boost model, demonstrating the ability to identify the discriminative pattern in the sEMG feature set. The lowest accuracy is obtained for the Cat Boost model, with an accuracy of 37%. While the model Multilayer perceptron and Gradient Boost obtain the accuracy of 66.17% and 64.14%, respectively. The findings combine deep learning with several boosting methods for efficient classification of phonetic language.

Keywords: Orthography, Surface electromyography(sEMG), Multilayer perceptron, Boosting algorithms.

INTRODUCTION

Language is the primary medium through which humans communicate ideas, thoughts and emotions. It mainly consists of an organised set of sounds, symbols defined by specific rules that allow meaningful interaction. In spoken language, speech production involves the coordinated activity of the vocal tract, articulatory organs, and facial muscles, whereas written language depends on orthographic representation to encode linguistic information [1].

Orthography is the system of writing a specific language. It mainly deals with the interconnection of spoken words and written graphemes [2]. Based on this, languages can be broadly categorised into phonetic and non-phonetic of origin.